Get the following:

Program: CORR

Groups: AIRFARE
BLKS
FUEL
GESSELL
HYPOCORR
PIKETOLL
RESID
ROOKIES
TVLIFE

CHAPTER 7: Scatterplots and Correlation

- Shows the relationship between 2 _QUANTITATIVE___ variables
- Can show categorical variables by _COLOR OR SYMBOL_____
- Individuals are represented by the ___DOTS___on the plot
  - **duplicates
- Explanatory Variable: (independent variable) *controlled by researcher
  - On the __X__ axis
  - Explains or causes the change in the __Y__ variable
- Response Variable: (dependent variable) try to predict
  - On the __Y__ axis
  - Measures the outcome of an experiment or study

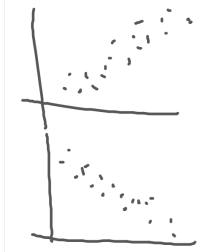
Interpreting Scatterplots:

Overall pattern

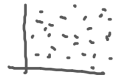
- Get a sense of what the data/plot looks like in general, then ***comment on the following 3 things*** Form, Direction, Strength

(1) Form

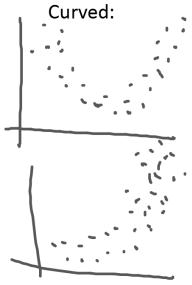
Linear:



Scattered



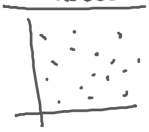
Curved:



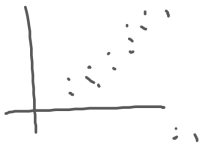
(2) Direction

Negative Association

No assoc.



Positive Association



(3) Strength- How well the points fit a form (curved or linear)

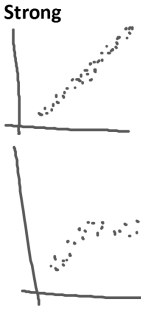
- Use the following words (or combinations of these):

- weak
- moderate
- strong
- scattered

mod. weak — mod. strong

Examples:

Strong



Moderate

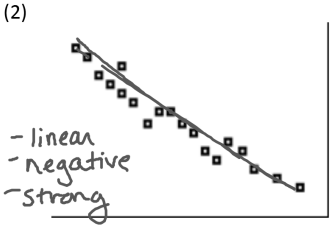
Weak

Scattered

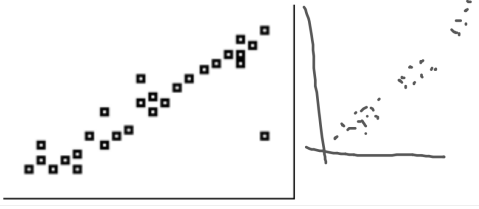
Examples:

(1) curved
- pos. assoc.
- strong

(2)

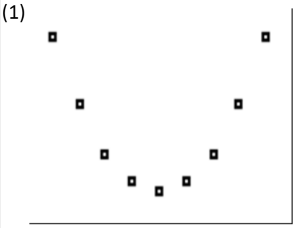


(3)
- linear
- pos.
- mod. strong
- outlier @ (x, y)



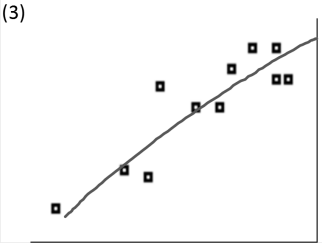
You try these:

(1)

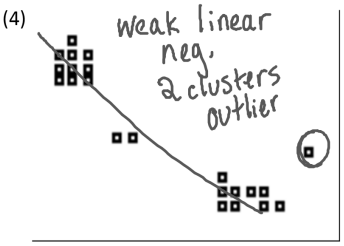


(2)

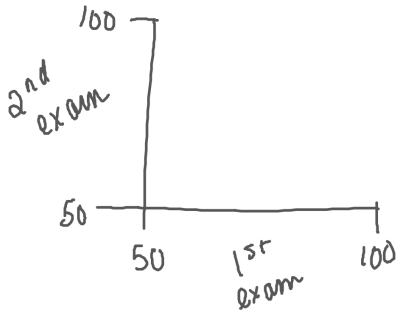
(3)



(4)



**** Now complete Worksheet 7A about describing scatterplots**



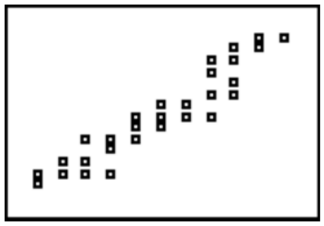
A. linear, positive, moderate
B. linear, weak, positive
C. linear, negative, strong
D. linear, negative, moderate
E. linear, positive, strong
F. linear, negative, weak, possibe outlier

CHART:

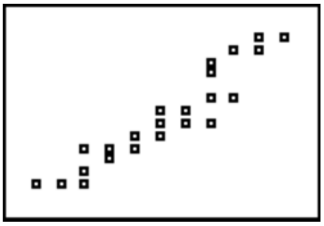
C	D	F
E	A	B

WHAT PLOT HAS THE STRONGER RELATIONSHIP? BY HOW MUCH??

A



B



Correlation

Symbol: $r = \#$

Definition:

Measures the strength and direction of the linear relationship between X and Y variables.

Formula:

$$r = \frac{1}{n-1} \sum \left(\frac{x_i - \bar{x}}{s_x} \right) \left(\frac{y_i - \bar{y}}{s_y} \right)$$

z_x z_y

no units

* Grouping/Ungrouping- GESELL
  2nd --> MEM --> #8: GROUP --> UNGROUP

* How to make scaterplots on the calculator
  STATPLOT--> SCATTERPLOT --> Put in X and Y lists --> ZOOM 9

* Deleting a point from a list

* Using the program CORR

****Complete worksheet 7B**

$\bar{x} = 14.381$
$\bar{y} =$
$s_x =$
$s_y =$
$r = -0.6403$

Correlation coefficient..... the last few notes.....

* $-1 \leq r \leq 1$
* close to 1 or -1 => strong linear relationship
* close to 0 => weak linear relationship
* - = negative association
  + = positive association
* -1 or 1 => perfectly straight line
* correlation is the same no matter which variable is X or Y
* correlation is the same even if you change the units that the variables are measured in (or manipulate the units)
  ***exception: changing sign, like multiplying by -3*
* correlation is NON-RESISTANT. It is affected by outliers
* correlation has NO units. It is just a number.

** BE CAREFUL!! Correlation does not mean causation!

r = 0.87 ==> strong, positive ASSOCIATION

this does NOT mean that the X variable causes the Y variable to change

READ p. 157 and p. 160 -- 161